Package 'genCountR'

May 8, 2026

Title Interacting with Roberts and Utych's (2019) Gendered Language Dictionary

Version 1.0.0

Description Allows users to generate a gendered language score according to the gendered language dictionary in Roberts and Utych (2019) <doi:10.1177/1065912919874883>.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.2.3

Depends R (>= 2.10)

LazyData true

Suggests devtools, knitr, rmarkdown, testthat

VignetteBuilder knitr

URL https://gencounter.app.damoncroberts.com, https://damoncharlesroberts.github.io/genCountR/

NeedsCompilation no

Author Damon Roberts [aut, cre] (ORCID: <https://orcid.org/0000-0002-4360-3675>)

Maintainer Damon Roberts <damon.charles.roberts@gmail.com>

Repository CRAN

Date/Publication 2023-12-05 17:50:04 UTC

Contents

dict — *Data from Gendered Language Dictionary Developed by Roberts and Utych (2019)*

Description

Each word in the dataset contains a rating by human coders. See details of dataset in the original paper.

Usage

```
dict
```

Format

A tibble with 701 rows and 15 columns:

Word The word to match

POS Part Of Speech

mean.a Mean score provided by all participants

std.dev.a Standard deviation of score provided by all participants

Source

https://journals.sagepub.com/doi/10.1177/1065912919874883

gen_count — *gen_count*

Description

Counts the number of masculine and feminine words in the document

Usage

```
gen_count(text)
```

Arguments

text	(string): A string object.

Details

Takes the number of words that are loosely categorized as Masculine, Feminine, or Neutral based on Roberts and Utych's (2019) definition. Feminine words had a score below 2.5, Neutral words had a score higher than 2.5 and lower than 5.5, Masculine words had a score higher than 5.5.

Value

data.frame with each word from the dictionary matched with the text and its number of occurances.

Examples

```
text <- 'This person was a heroine due to their fighting during the war.'
result_df <- genCountR::gen_count(text)
```

---

gen_score *gen_score*

Description

Calculates the score of the supplied text string based on the Gendered Language Dictionary created by Roberts and Utych (2019).

Usage

```
gen_score(text)
```

Arguments

text	(string): A string object

Details

Takes the matched words and their occurrence in the supplied text, finds the score for those matched words in the dictionary, sums those scores up and then divides it by the total number of words in the dictionary.

Value

list object with avg_score of the supplied text string, total_score of supplied text string, and data.frame of matches

Examples

```
text <- 'Hero. hero Heroine. heroine, Prison. Prisom.'
result <- genCountR::gen_score(text)
result$avg_score
result$total_score
result$df
```

text_clean *text_clean*

Description

Cleans the supplied text string and converts it into a list of individual words.

Usage

```
text_clean(text)
```

Arguments

text	(string): A string object.

Details

Takes the string, converts all the characters to lower case, removes punctuation, and splits the string into individual words.

Value

list of each word in all lower case and without punctuation.

word_count *word_count*

Description

Count number of times a particular word from the dictionary shows up in a document.

Usage

```
word_count(word_item)
```

Arguments

word_item	(vector or array): A vector or array of words from the text wanting to be matched to dictionary.

Details

Takes the number of words that are loosely categorized as Masculine, Feminine, or Neutral based on Roberts and Utych's (2019) definition. Feminine words had a score below 2.5, Neutral words had a score higher than 2.5 and lower than 5.5, Masculine words had a score higher than 5.5.

Value

data.frame object of with count of masculine, feminine, and masculine words.

Index